

## Listing of Claims

1-30. (Canceled)

31. (Currently Amended) A method for linking grammars into a hierarchy, comprising the operations of:

establishing various grammars each grammar including various utterances and, for each utterance, the following associated attributes: (1) an indication of whether the utterance is explicitly [[linked]] chained to a further grammar, or (2) contextual information [[metadata]] indicating a type of data implicitly specified by the utterance;

where each one of the various grammars further includes, for each utterance that is explicitly chained [[linked]] to a further grammar, a chained command attribute [[link]] identifying the further grammar for activating [[importing]] responsive to a user issuing that utterance while said further grammar is activated for speech recognition;

where the various grammars include command grammars and information-type grammars, and:

utterances in the command grammars form commands to control a manner of presenting video programs;

utterances in the information-type grammars form keywords pertaining to content of video programs;

accepting a series of user utterances; and

performing a series of operations by activating further grammars, wherein for each user utterance that is explicitly chained to a further grammar, activating the further grammar; and for each user utterance having contextual information associated within, activating a further grammar based on the contextual information of a preceding grammar.

32. (Previously Presented) The method of claim 31, further comprising operations responsive to receiving an utterance in a given information-type grammar while said given grammar is activated for speech recognition, comprising:

if said given grammar lacks a link from the utterance to a further grammar, processing said utterance based on (1) application context of a user-driven system for presenting video programs and (2) type of data specified by the received utterance according to the given grammar.

33. (Currently Amended) The method of claim ~~[[30]]~~ 31, the type of data indicated by the contextual information ~~[[metadata]]~~ in information-type grammars is selected from a list that includes:

program name, genre, actor, director, writer, episode, date, popularity, quality rating, subject matter rating.

34. (New) The method of claim 31, wherein the utterances comprise user defined preference settings.

35. (New) The method of claim 34, wherein the user defined preference settings are selected from among a subset of program categories, a popularity requirement, a parental-warning type rating, and a quality requirement.

36. (New) The method of claim 31, further comprising:

deducing predicted preferences from the various utterances, wherein the predicted preference settings are defined by processes selected from among a viewing pattern analysis, a user profile analysis, an analysis of user behavior relating to frequency of content requests by way of utterance.

37. (New) The method of claim 36, wherein the predicted preference settings are added to a user defined preference setting.

38. (New) A method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source comprising:

- assembling a hierarchical collection of operational grammars;
- providing a set of various grammars comprising command grammars and information grammars, wherein the various grammars include various utterances and various attributes, the various attributes comprising information relating to the utterances contained within a grammar;
- receiving a first utterance from a user;
- activating at least one first grammar containing at least part of the first utterance, wherein the at least one first grammar matches the first utterance to a command grammar or an information grammar, or both, forming a first operation instruction;
- traversing the hierarchal collection of operational grammars and applying a first operation to a first utterance;
- preparing the speech recognition system to receive at least one additional utterance from a user by activating at least one additional grammar based on the various attributes contained within the at least one first grammar;
- receiving at least one additional utterance from the user;
- activating at least one additional grammar containing at least part of the additional utterance, wherein the at least one additional grammar matches the additional utterance to a command grammar or an information grammar, or both, forming at least one additional operation; and
- progressively refining a set of operations by traversing the hierarchal collection of operational grammars and applying at least one additional operation.

39. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source

according to claim 38, further comprising deactivating the at least one first grammar before activating the at least one additional grammar.

40. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 38, wherein the various attributes of the various grammars at least partially include explicit links to at least one additional grammar, the method further comprising:

explicitly linking the at least one first grammar to at least one additional grammar based on an explicit link contained within the at least one first grammar.

41. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 38, wherein the various attributes comprise user defined preferences.

42. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 41, wherein the user defined preferences are selected from among a subset of program categories, a popularity requirement, a parental-warning type rating, and a quality requirement.

43. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 38, wherein the various attributes comprise predicted preferences, wherein the predicted preferences are defined by processes selected from among a viewing pattern analysis, a user profile analysis, an

analysis of user behavior relating to frequency of content requests by way of utterance.

44. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 41, wherein the step of activating at least one additional grammar further comprises:

- activating more than one additional grammar;
- ranking the more than one additional grammar based on the presence of preferences within the more than one additional grammar; and
- presenting a user with options for the at least one additional utterance based on the ranking of the more than one additional grammar.

45. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 43, wherein the predicted preference settings are added to a user defined preference setting.

46. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 38, wherein more than one first grammar is activated after receiving a first utterance from a user.

47. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 38, wherein more than one additional utterance is received

from the user, wherein more than one additional grammar is activated, forming more than one additional operation.

48. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 38, wherein utterances received from a user are analyzed for various traits, and wherein the various traits are used for targeted advertising.

49. (New) The method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source according to claim 48, wherein the various traits are selected from among utterance syntax, speaker identification, language identification, dialect identification and emotional state identification.

50. (New) A method of creating a speech user interface using a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, to construct an operational command hierarchy for controlling a video source comprising:

providing a set of various grammars comprising command grammars and information grammars, wherein the various grammars include various utterances and various attributes, the various attributes comprising information relating to the utterances contained within a grammar;

receiving an utterance from a user comprising at least a first portion and a subsequent portion;

activating at least one first grammar containing a first portion of the first utterance, wherein the at least one first grammar matches the first portion of the utterance to a command grammar or an information grammar, or both, forming a first operation instruction;

activating the at least one additional grammar containing a subsequent portion of the first utterance, wherein the additional grammar matches the subsequent portion of the utterance to a command grammar or an information grammar, or both, forming an additional operation instruction;

constructing a hierarchy of operational instructions by adding the first operational instruction to the hierarchy of operational instructions and adding at least one additional operation instruction to the hierarchy of operational instructions.

51. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 50, further comprising:

activating another additional grammar based on one or more various attributes included in the at least one first grammar or the at least one additional grammar.

52. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 50, further comprising:

deactivating the at least one first grammar before activating the at least one additional grammar.

53. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 51, wherein the various attributes comprise user defined preference settings.

54. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 53, wherein the user defined preference settings are selected from among a subset of program categories, a popularity requirement, a parental-warning type rating, and a quality requirement.

55. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 51, wherein the system response to a search or content selection command may depend in part upon predicted preference settings, wherein the predicted preference settings are defined by processes selected from among a viewing pattern analysis, a user profile analysis, an analysis of user behavior relating to frequency of content requests by way of utterance.

56. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 55, wherein the predicted preference settings are added to a user defined preference setting.

57. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 50, wherein utterances received from a user are analyzed for various traits, and wherein the various traits are used to select appropriate advertising messages to display to said user.

58. (New) The method of constructing a hierarchy of operational commands in a speech recognition system for controlling a video source according to claim 57, wherein the various traits are selected from among utterance syntax, speaker identification, language identification, dialect identification and emotional state identification.

59. (New) A method of managing a hierarchically organized collection of grammars, or equivalent single hierarchical grammar, in a speech recognition system to support multiple schemes of designing a speech user interface for controlling a video source comprising:

providing a collection of grammars comprising command grammars and information grammars, wherein the grammars include various utterances and various attributes, wherein the various attributes comprise information relating to the utterances contained within the grammars;

providing a first means for designing a speech user interface using a chained command string comprising:

assembling a hierarchical collection of operation grammars;

accepting a string of utterances from a user;

sequentially activating at least one command grammars and at least one information grammar based on the string of utterances, forming operational instructions; and

traversing the hierarchal collection of operation grammars and applying a first operation to a first utterance;

preparing the speech recognition system to receive at least one additional utterance from a user by activating at least one additional grammar based on the various attributes contained within the at least one first grammar;

receiving at least one additional utterance from the user;

activating at least one additional grammar containing at least part of the additional utterance, wherein the at least one additional grammar matches the additional utterance to a command grammar or an information grammar, or both, forming at least one additional operation; and

progressively refining a set of operations by traversing the hierarchal collection of operation grammars and applying at least one additional operation; and

providing a second means for designing a speech user interface using a multi-step string of commands comprising:

receiving a first utterance from a user;

activating a first speech grammar based on the first utterance that at least partially contains the first utterance, forming a first instruction;

adding the first instruction to an operational hierarchy;  
preparing the speech recognition system to receive at least one additional utterance from a user by activating at least one additional grammar based on the various attributes contained within the at least one first grammar;  
receiving at least one additional utterance from the user;  
forming at least one additional instruction; and  
refining the operational command hierarchy by adding the at least one additional instruction to the operational command hierarchy;  
receiving a first utterance from a user wanting to control a video source;  
determining a means for designing a speech user interface comprising:  
analyzing the first utterance and selecting from among the first means and the second means, wherein the first means is selected when the first utterance indicates that the utterance is a chained command, and wherein the second means is selected when the first utterance indicates that the utterance is a first portion of a multi-step set of commands.